REPRESENTING MENTAL MAPS AND COGNITIVE LEARNING IN MICRO-SIMULATION MODELS OF ACTIVITY-TRAVEL CHOICE DYNAMICS

T.A. Arentze and H.J.P. Timmermans, Eindhoven University of Technology

Conference paper
Session XXX

Moving through nets:
The physical and social dimensions of travel
10th International Conference on Travel Behaviour Research

Lucerne, 10-15. August 2003

REPRESENTING MENTAL MAPS AND COGNITIVE LEARNING IN MICRO-SIMULATION MODELS OF ACTIVITY-TRAVEL CHOICE DYNAMICS

T.A. Arentze and H.J.P. Timmermans

Eindhoven University of Technology

Urban Planning Group

PO Box 513, 5600 MB Eindhoven, The Netherlands

Phone: +31 40 247 3315
Fax: +31 40 243 8488
eMail: m.v.d.sande@bwk.tue.nl

Abstract

This paper develops a model, based on Bayesian beliefs networks, for representing mental maps and cognitive learning into micro-simulation models of activity-travel behavior. Mental maps can be used to address the problem that choice sets in models of travel demand are often ad hoc specified. The theory underlying the model is discussed, a specification is derived and numerical simulation is used to illustrate the properties of the model. The model reported in this paper is part of a wider research effort to model various aspects of learning and adaptation behavior in urban settings. These models will ultimately linked to the Albatross model system.

Keywords

Mental maps, activity-travel behavior, learning and adaptation behavior, Bayesian belief networks, micro-simulation models

Preferred citation

Arentze, Theo and Harry Timmermans (2003) Representing Mental Maps and Cognitive Learning in Micro-Simulation Models of Activity-Travel Choice Dynamics, paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, August 2003.

1. Introduction

Transportation science has a long tradition in developing and applying choice models to predict transport mode, destination and route choice. Traditionally, many models have been developed to predict single choice facets. More recently, models predicting multiple choice facets have been proposed, culminating in complex activity-based models of travel demand.

Although the theoretical underpinnings of these models may differ, they have in common the assumption that individuals will choose the alternative *within their choice set* they prefer, sometimes subject to some constraints. In the vast majority of these models, however, the construction and composition of individual choice sets is not explicitly modeled. Choice sets are typically arbitrarily assumed, or derived on an ad hoc basis by the researchers using some arbitrary rule (e.g. a travel time band).

As long as the IIA-property is satisfied, the composition of the choice set has no implications for the estimation of the utility function. However, choice set composition will affect predicted market shares as the latent demand is allocated to the alternatives belonging to an individuals' choice set. Thus, predictions of market shares of choice alternatives will be biased if individual choice sets are misspecified. If the composition of the choice set also influences individual preferences and choice behavior, the parameters of the estimated utility or preference function will also be biased.

The literature on environmental cognition (e.g., Golledge, 1993; Horton and Reynolds, 1971; Smith, 1976; Potter, 1979; van der Heijden and Timmermans, 1982; Timmermans, et al, 1984) suggests that people have limited information about their environment. They are not necessarily familiar with all the choice alternatives in their environment. Individuals learn about their environment during the implementation of their activities. Repeated choices make them (better) aware of some choice options, which may induce them to consider a destination, choose a route or try a new transport mode. Thus, the implementation of activities leads to dynamics in the mental representation of choice alternatives, which constitutes the basis for choice.

In addition to the above mechanisms, individuals may also decide to become involved in active spatial search. For example, a move to a new unknown city or area implies the need to explore the area, try different alternatives and in doing so build up a choice set. Likewise, negative experiences with existing alternatives may prompt individuals to actively search and try alternatives, again leading to changes in existing choice sets. Spatial search, cognitive learning and dynamic choice sets are all ignored in most existing models of spatial choice. This lack of attention can be reduced to a fundamental flaw, namely: mental maps of individuals are not explicitly represented in the models.

The purpose of the present study therefore is to develop a model of mental maps and show how this can be integrated in discrete choice models/activity-based model to simulate dynamic decision-making under uncertainty, spatial search and spatial cognitive learning. The results envisioned are particularly relevant for micro-simulation of activity-travel choice in space and time. A mental map is defined as an individual's mental representation of his/her environment and includes the beliefs and values about the attributes characterizing the alternatives in the environment. Hence, an individual will typically have imperfect and incomplete information about his/her environment. A full representation of a mental map will include all these aspects. In the present paper, however, we will focus on the perceptual and cognitive aspect. Hence, an individual will typically have imperfect and incomplete information about his/her environment. We represent the mental map of individuals as a Bayesian Belief Network (BBN) and model cognitive learning as updating of beliefs in the network in response to observations the individual makes when (s)he implements trips and activities. We argue that this approach has several potential advantages. First, learning is incremental so that adaptive behavior can be modeled in a natural way. Secondly, beliefs are represented as probability distributions so that they can be integrated in a utility framework to model decision-making in a straightforward way (namely through the concept of *expected* utilities). Thirdly, beliefs are represented as probability distributions so that the degree of uncertainty and expected information gain can be quantified by means of an entropy measure.

2. Theory

2.1 Definition of the problem

Consider an individual who just moved to a new area and therefore has limited knowledge about the spatial environment in which (s)he lives and implements his/her daily activities. Having limited knowledge, the attributes of locations/destinations and routes are generally uncertain. During trips and activities, the individual makes observations on locations and links of the network that allow him/her to update his/her beliefs and increase his/her knowledge about the area. Observations are not necessarily perfect and the individual takes into account possible error in updating beliefs meaning that there may still be uncertainty left. Furthermore, (s)he uses knowledge about how urban areas are structured in general and derives expectations if specific information about a location has become available such as for example the type of urban area in which it is located (e.g., inside or outside the inner city), the location's position relative to the road network, the land use in the neighborhood and so on.

The present study addresses the problem of how mental maps and spatial learning can be modeled and integrated into micro-simulations of activity-travel choice. We will focus here on the part of mental maps that is concerned with the spatial environment and conveniently assume complete knowledge about the transport network. Furthermore, we will concentrate on static attributes such as physical characteristics of locations and leave more dynamic attributes such as (congestion-dependent) waiting times out of consideration. These problems are addressed in other ongoing projects, and the results of these projects will be combined at a later stage.

2.2 Formalization

Although it is not critical for the approach proposed here, assume that the environment is represented as a regular grid of cells. Each cell represents a location where possibly an activity can be conducted and is described by a vector, X_l, of potentially relevant variables for location choice. A transport network connects the different locations and is modeled, as usual, by a directed graph $G(N, L)$ where N is a set of nodes and L is a set of links. A trip from an origin location l_1 to a destination location l_2 is modeled by a path through the network from the nearest node from l_1 to the nearest node from l_2. The variables that describe locations are considered to be discrete or discretisized. The possible values (or states) of X_{lk} are denoted by x_{lks}, where x_{lks} is a specific value (or 'state') of X_{lk} and l is an index of location. The belief that $X_{lk} = x_{lks}$ is represented by a probability, $P(x_{lks})$. Full information or certainty is represented as a special case, namely the case where the probability is zero for untrue values of the variable and one for the true value. Thus, the mental map of an individual is modeled by a set of probability distributions, $\Pi = \{P(X_{lk}) \mid l = (i, j), i = 1,\ldots, I, j = 1,\ldots, J, k = 1,\ldots, K\}$ where I and J are the number of rows and columns of the grid and K is the number of location attributes. Individuals are represented as agents that schedule their activities and execute their activity schedules in space and time. Consequently, at any moment in time an individual is either conducting an activity in a cell or is traveling on a link of the network. In the course of both types of actions, the individual may make observations that change his/her beliefs in Π. Therefore, each time after completing an activity or trip, the system simulates possible observations made during the activity or trip and updates Π according to the outcome of the observation.

2.3 Observations and updating of beliefs

Observations can be conceptualized as communication (between the observed and observer) over a noisy channel. Assuming that the subject (observer) is aware of possible error, we propose a Bayesian method of belief updating based on observations as follows. The unit of observation is a certain variable X in a certain cell. Let Y denote the outcome of the observation, whereby $Y = y_s$ denotes the outcome that $X = x_s$ in that cell. Possible error of observations means that the probability of $Y = y_s$ given that $X = x_s$ is not necessarily one and the probability that $Y = y_s$ given that $X \neq x_s$ is not necessarily zero. Given some observation outcome y_u, the belief in $X = x_s$ is updated according to the Bayesian method as:

$$P(x_s \mid y_u) = \frac{P(y_u \mid x_s)P(x_s)}{\sum_{s'=1}^{n} P(y_u \mid x_{s'})P(x_{s'})} \quad \forall s \tag{1}$$

where:

$P(x_s)$ is the prior belief in x_s;

$P(x_s \mid y_u)$ is the updated belief after observation y_u;

$P(y_u \mid x_s)$ is the probability of observation $Y = y_u$ given $X = x_s$;

n is the number of possible values of X and Y.

The updated belief, $P(x_s \mid y_u)$, is the prior belief in a next observation so that learning indeed is incremental.

Equation (1) assumes that the conditional probability table, $P(Y \mid X)$, is known and taken into account by the individual in updating beliefs. The table defines the observation-outcome probabilities under each assumption of the actual value of the variable observed. Thus, the method assumes that the subject is aware of a probability of making errors in observations and takes this probability into account in updating his/her beliefs. In the extreme case where the observation is completely insensitive, the cells in the table would equal:

$$P(y_u \mid x_s) = 1/n \qquad \forall u, s \tag{2}$$

where n is the number of possible values of Y (and X). Effectively, this limiting case is identical to making no observation at all, because it implies a uniform observation-outcome probability distribution for each value of X. It is easy to see that, in this case, according to (1) the updated beliefs stay equal to the prior beliefs, meaning that the observation has no impact on beliefs. In the other extreme case, the observation is maximally sensitive to values of X and we would have:

$$P(y_u \mid x_u) = 1 \text{ and } P(y_u \mid x_s) = 0, \;\; \forall s \neq u, \forall u \tag{3}$$

It is easy to see that, in this case, according to (1) the updated beliefs equal $P(x_u \mid y_u) = 1$, and $P(x_s \mid y_u) = 0$, $\forall s \neq u$, if the outcome is y_u. In other words, in this limiting case, any observation would reduce the uncertainty completely. In all other cases, where a subject is aware of the probability of making errors, we have:

$$1/n < P(y_u \mid x_u) < 1 \text{ and } 0 < P(y_u \mid x_s) < 1/n, \forall s \neq u, \forall u \tag{4}$$

It is easy to see from (1) that, in this case, $P(x_u \mid y_u) > P(x_u)$ and $P(x_s \mid y_u) < P(x_s)$, $\forall\ s \neq u$, meaning that an observation increases the belief in the outcome of the observation and decreases the belief in all other values of the variable.

To use equation 1 for belief updating, we need a function to predict $P(Y \mid X)$ in any possible state of the individual and the system and for each observational variable. To derive such a function, we assume that observation accuracy is composed of two factors, namely a sensitivity and bias. Sensitivity is conceptualized here as the inverse of the amount of error on the scale of which equations (3) and (4) constitute the upper and lower extreme, respectively, and bias is conceptualized as the likelihood of confusing values x_s and x_u after correcting for the error scale. Therefore, we propose to use a logit model defined as follows:

$$P(y_u \mid x_s) = \frac{\exp(\theta\beta_{us})}{\sum_{s'=1}^{n} \exp(\theta\beta_{us'})} \qquad \forall u,s \tag{5}$$

where β_{us} are observation-bias parameters and θ is the observation-sensitivity parameter. Keeping sensitivity, θ, constant, an increase in β_{us} leads to a higher probability of confusing u for s. Observation bias would be absent if, for all u, $\beta_{uu} > 0$ and $\beta_{us} = 0$, $\forall s \neq u$. On the other hand, scale parameter, θ, meets the conceptual requirements expressed by equations (2) – (4). For $\theta = 0$ we find $P(y_u \mid x_s) = 1/n$, so that condition (2) is met and for increasing values of θ, $P(y_u \mid x_u)$ goes to 1 and $P(y_u \mid x_s)$ goes to 0, $s \neq u$, so that also conditions (3) and (4) are met.

Having reduced table $P(Y \mid X)$ to a scaled logit function, the next question becomes how the beta and scale parameters involved can be set to simulate behavior. Regarding the beta parameters, some points on the scale can be identified logically. Since θ sets the scale of the parameters, it is most natural to use a zero-one scale for the bias parameters, whereby $\beta_{uu} = 1$, $\forall u$, and $\beta_{us} = 0$ for values x_u and x_s that are considered to be most easily confused with each other for the variable considered. On the other hand, we assume that the sensitivity parameter is a function of the following factors known in the system:

$$\theta = f(E, X, D, M, A, L) \tag{6}$$

where E is the type of event (activity or trip), X is the type of variable the observation relates to, D is distance from the cell when the observation is made, M is transport mode (in case of a trip), A is the activity type or trip purpose depending on the event and L is the type of link (in case of a trip). To give some examples, sensitivity is higher if the variable type is easy to observe, the distance from the cell is small, transport mode is slow and the link type is local road (as opposed to highway). The activity type or purpose of the trip is included for its possible impact on the individual's motivational state. For example, when the purpose of the trip is information seeking observation sensitivity is probably higher than on a routine trip from work to home. Also, interactions between X and A may be influential in the sense that certain trip purposes (e.g., shopping) increase or decrease the sensitivity to certain variables (e.g., presence of stores).

Updating the mental-map Π of a simulated individual involves, for each cell l and each variable k, calculating conditional probabilities of observation outcomes based on equation (5), simulating an observation and updating the current beliefs in Π using equation (1). Simulating an observation is done by Monte Carlo drawing from the probability distribution $P(Y \mid x_s)$, where x_s is the true value (of variable k in cell l). One of the advantages of including sensitivity as a continuous parameter of observations is that the system does not need to decide whether or not an observation is made. Not making an observation is simply a special case of making an observation, namely one with zero sensitivity. As long as the sensitivity function (equation 6) is properly specified, the deductions should be adequate.

2.4 Representing mental maps as a Bayesian Belief Network

The mental map is not considered as a collection of isolated beliefs, but rather as a set of beliefs that are interconnected. Links between beliefs represent causal or statistical relationships between variables that allow the subject to make inferences about, in this application, the land-use and transport system. Due to such inferences the evidence obtained for one belief tends to spread across the network and, therefore, has consequences for other beliefs as well. Given that we use the Bayesian method for updating beliefs, the method that we propose here qualifies as a Bayesian belief network. This kind of network has been and still is intensively studied in areas of Artificial Intelligence, Statistics, Decision Analysis and Operation Research for application in probabilistic expert systems (Heckerman et al. 1995, Russel and Norvig 1995, Spiegelhalter et al. 1993). At present, efficient algorithms are available for belief updating in the context of such networks.

A Bayesian belief network defines a causal structure between beliefs by a directed, a-cyclic graph. Developing a belief network model for the present purpose involves identifying the attribute variables, X_k, and the relationships between the variables that adequately represent the spatial knowledge that individuals use in activity travel choice. Consider as an example the simple specification of a belief network (Figure 1) that is used in the simulations described in the next section. This belief network includes the variables urban-area type (inner city, high urban density, low urban density and outer area), land use (industry, housing, commercial, green, mixed, other), availability of facilities for specific activities (yes, no) and attractiveness of locations for specific activities (zero, low, medium, large). Availability and attractiveness are included for only three activities (namely shopping, leisure and recreational activities). However, it is easy to see how the network can be extended to cover other activities as well. The arcs represent the supposed causal relationships. For example, expectations about the

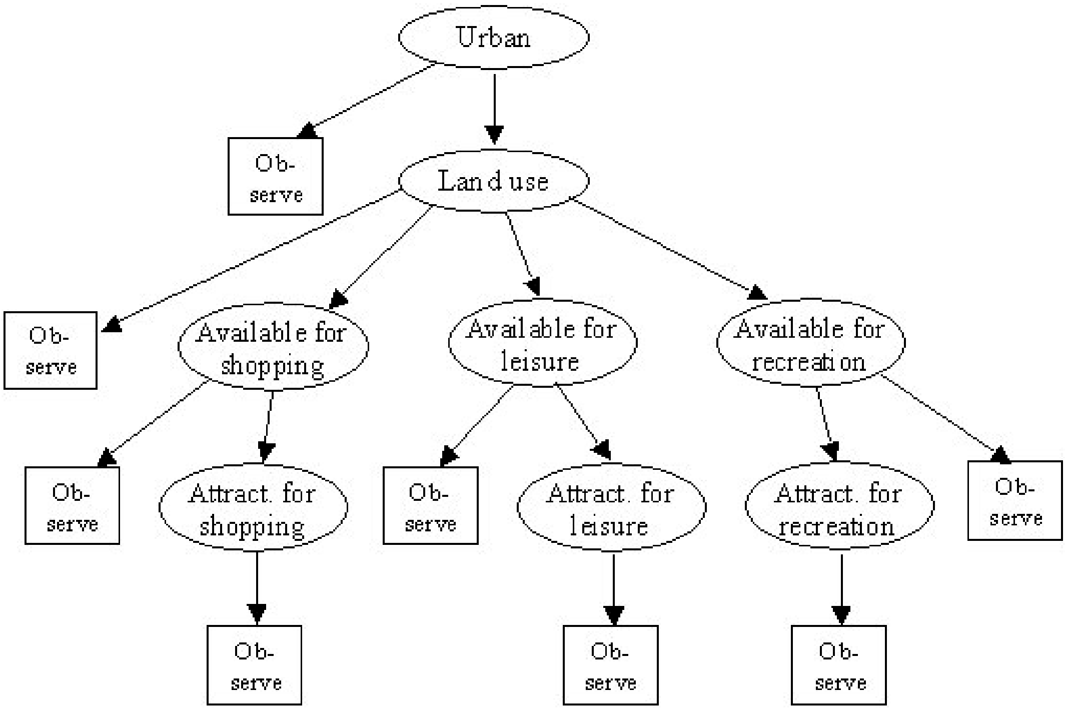


Figure 1. The Bayesian-belief-network model used in the simulations

land use of a location are influenced by urban-area type and expectations about availability of specific facilities are dependent on the land use. In this particular case, the graph of the network is a tree. This is, however, not a formal restriction, as cross relationships can be included as well as long as the graph maintains a-cyclic.

Attached to each node is a so-called node probability table (NPT) defining a-priori beliefs ($P(Y \mid R)$) for the node (Y) under each possible combination of states (r_s) of the parent variables. For root nodes, which do not have parents, the NPT reduces to an unconditional probability distribution. To give an example, Table 1 displays an arbitrary NPT for the node representing the conditional beliefs for the availability of shopping facilities. In general, the structure and NPTs of a network represent (generic) knowledge of an individual. This knowledge may not be constant across persons and spatial contexts. Persons may hold different beliefs about how urban areas are structured depending on their specific experiences and level of expertise. In applications we suggest a compromise solution in which the network is taken as given and fixed and the included NPTs are estimated on data.

Table 1. Example: arbitrary NPT related to availability of shopping facilities

	Shopping facilities (Pr)	
Land use	Yes	No
Industry	0	1
Housing	0.20	0.80
Commercial	0.50	0.50
Green	0	1
Mixed housing-commercial	0.33	0.67
Other	0	1

In addition to the variables that have substantial meaning, the network includes observation-outcome variables. The arcs run from substantial to observational nodes and not vice versa, to reflect the fact that the values of substantial variables affect observation-outcome probabilities and not vice versa. Moreover, connected in this way the NPT of each observation node has the structure $P(Y \mid X)$ and, hence, is consistent with the method of belief updating outlined in the previous section (Y is the outcome and X is the object of the observation). The observation nodes are the (only) entries through which new information enters the network. An observation is modeled as hard evidence for the observation node. Evidence spreads through the network by backward and forward reasoning. Backward reasoning involves applying the Bayesian method given by equation (1) to update beliefs related to parents given evidence entered to child nodes. In forward reasoning the beliefs of child nodes are adjusted to make them consistent with changed beliefs of parent nodes. Efficient algorithms for updating Bayesian belief networks based on these (Bayesian) principles exist. Using the updating algorithm means that at any moment in time all beliefs are consistent with evidence, given by observations, and spatial knowledge, given by the NPTs.

Several levels of knowledge can be modeled by NPTs as follows. First, the complete absence of causal knowledge is modeled by repeating the same probability distribution in each row of the table so that in effect all beliefs are unconditional. Under this setting, the individual still learns from observations, but beliefs are not interconnected so that learning is slower (and unbiased). Second, the individual may know the (typical) marginals of distributions (e.g., land use or facilities) and hence the a-priori probability of finding a particular state of the variable in a cell. This knowledge puts a constraint on the NPTs as follows:

$$P(x_u) = p_u^x \qquad \forall u \tag{7}$$

for all root nodes X and

$$\sum_{s=1}^{S} P(y_u \mid r_s) P(r_s) = p_u^y \qquad \forall u \tag{8}$$

for all other nodes Y, where r_s is a combined state of parent nodes, S is the number of combined states of parent nodes, y_u is the u-th level of the node and p_u^x and p_u^y are given, a-priori probabilities of X and Y. Hence, this knowledge level can be represented by adjusting $P(y_u \mid r_s)$ such that constraint (7) is met for each level y_u and each node Y (working from root nodes downwards in the directions of the arcs). Finally, the individual may have perfect (generic) knowledge meaning that the conditional probabilities perfectly correspond to the real conditional probabilities. To simulate that limiting case, we set the NPTs to the true values in the study area.

NPTs represent generic knowledge in the sense that they apply to all locations (cells). This is complementary to the *location-specific* probability distributions stored in the mental map that represents the individual's current beliefs related to specific locations. The complete absence of location-specific knowledge is modeled by setting the probability distributions for each location to the a-priori probability distributions, i.e. the probabilities computed based on NPTs. Each time observations are made related to some location l, the network (set of NPTs) is instantiated by l and updated based on the (hard) evidence entered at the observation nodes. Instantiating the network means adopting the current probability distributions for l as the prior probabilities at nodes. The updated probabilities replace the existing probabilities for l and constitute the new state of the mental map. In this way, evidence obtained by observations is accumulated and incorporated in the current beliefs for each location.

The difference between the current belief and the a-priori belief in a certain value for a specific cell represents the level of specific knowledge about that value in that cell. We assume that due to limited memory retention capacity, a certain decay of specific knowledge occurs in each time step in the system. As a consequence of the decay, beliefs return to their corresponding a-priory belief with a certain step size in each time step (cf. Arentze and Timmermans, 2003; Timmermans, Arentze and Ettema, 2003). In equation:

$$p_{lu}^{t} = p_u^0 + \alpha(p_{lu}^{t-1} - p_u^0) \tag{9}$$

Rewriting gives:

$$p_{lu}^{t} = (1-\alpha)p_{u}^{0} + \alpha p_{lu}^{t-1} \tag{10}$$

where ${p_{lu}}^t$ is the belief in value u for cell l at time t, ${p_u}^0$ is the a-priori belief and α is the step size parameter, whereby $\alpha = 1$ means no decay and $\alpha = 0$ means complete decay in every time step. Thus, the decay undoes the effects of observations meaning that if no observations are made for a sufficiently long time the belief will have been completely returned to the starting point. The step size parameter does not necessarily have the same value for each variable, as memory retention may differ depending on characteristics such as saliency and relevance of the variable.

3. Illustration

The model was implemented in a C++ program. For illustration purpose, a simple belief network was implemented (Figure 1) together with Pearl's method (Pearl 1988, see also Neapolitan 1990) for updating beliefs in a Bayesian Belief Network. The method of belief updating proposed by Pearl was designed as a model of how humans adjust their beliefs and, therefore, suits our present modeling purpose. In this method, nodes undergoing a change send messages to parent and child nodes. In the receiving node, a message from a child or parent triggers adjustments of existing beliefs so as to make them consistent with the new evidence. The method works optimally only for networks in which the graph is a tree, such as the network assumed here.

In this section we discuss results of simulations that were conducted to explore the behavior of the model and show how it can be used in micro-simulations of activity-travel patterns. The model predicts the learning path of an individual whose activities are simulated in time and space. First, we describe the data of the hypothetical case.

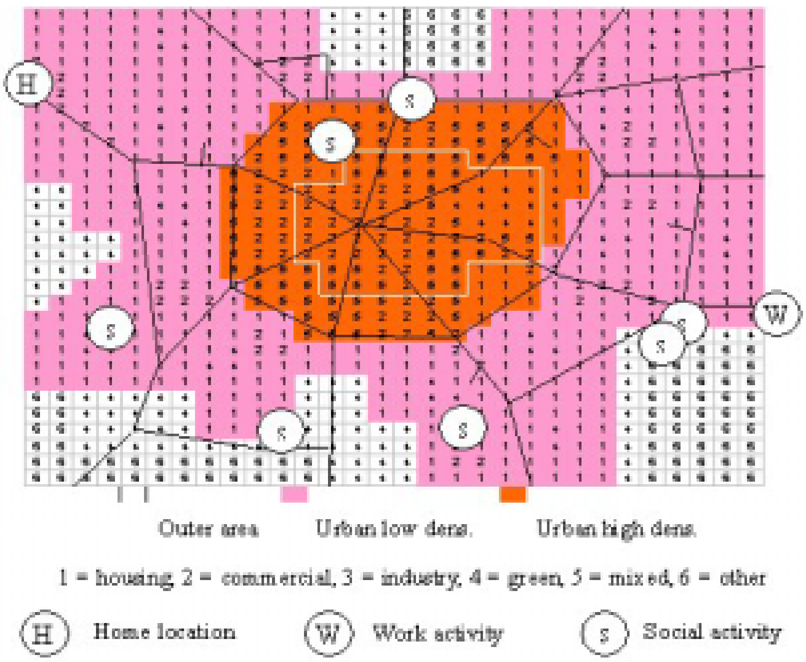


Figure 2: Map of the study area and fixed activity locations

3.1 Data of the case

The hypothetical study area is modeled by a 32-by-32 raster of cells. Figure 2 shows the assumed distributions of the urban-area-type variable (by shades of gray) and land-use variable (by numbers). The white lines indicate the borders of the inner city and the black lines represent the transport network. For the purpose of illustration, only a limited number of facilities were taken into consideration. These included shopping facilities, leisure (restaurants, cafes, etc.) and green recreation (parks). The circles show the residential and fixed activity locations of the hypothetical individual. The activities include work, social, shopping, leisure and recreational activities. The locations for work and social activities are taken as given, whereas the locations for the other activities are free to choose. All work activities take place at the same location. Since this location is located outside the study area, the work trips are monitored till the point where they leave/enter the study area. The location for social activities varies across specific episodes of the activity. For the destination choice of shopping, leisure and recreational trips, a simple decision rule is used. The rule selects the location maximizing the expected utility based on current beliefs about attractiveness. The utility function used includes parameters for distance and the (ordinal) attractiveness levels.

The transport network includes only main roads and consists of 54 links and 42 nodes. The data for each link includes a resistance (i.e., the length). All links can be traveled in both directions. Although the spatial and transport-network data are made up, they do convey a typical structure of urban areas (at least for Dutch cities). Figure 3 displays the distributions of the attractiveness variables related to the three activities.

Each activity includes a trip from home to the activity location, an activity at the trip destination and a return trip. A shortest-path algorithm maps the trips on the road network. After each trip/activity event, the system predicts an observation for each cell and each attribute and updates the beliefs accordingly. Basic settings were adopted for the observation-outcome probability function. All β_{us} were set to zero for off-diagonal cells ($u \neq s$) and to unity for the diagonal cells ($u = s$). This defines a case where there are no specific observation biases. The only exceptions are the attractiveness variables. For these variables complete confusion was assumed between non-zero levels (low, medium, high) reflecting the idea that attractiveness level cannot be observed from a distance. A basic specification was also assumed for the function used to predict thetas (equation 6), namely: $\theta = \max(0, 4 - 1.25D)$, where D is distance measured in cell widths and refers to the shortest (straight-line) distance across the links and nodes included in the trip. Thus, the function assumes that theta depends only on distance and is the same for all variables X_k. The thetas relate to observations made during the trip. For activity events it is assumed that for the cell in which the activity is conducted all variables are observed with certainty, whereas for all other cells no observations are made at all. Finally, the memory retention parameter, α, is set to one simulating the case were there is no decay of specific knowledge in time.

The initial state of the mental map assumes full information about the urban-area-type variable and complete absence of information about the other variables. Thus, for the latter variables initial probabilities are derived from NPTs. In turn, the specification of NPTs implies perfect general knowledge in the sense that the conditional probabilities all correspond to statistical data of the study area. In sum, the specifications are consistent with the case of an individual who just moved into a new city and knows nothing about available facilities, but has full information (e.g., through a map) of how the area is structured in terms of area type and has unbiased general knowledge about cities.

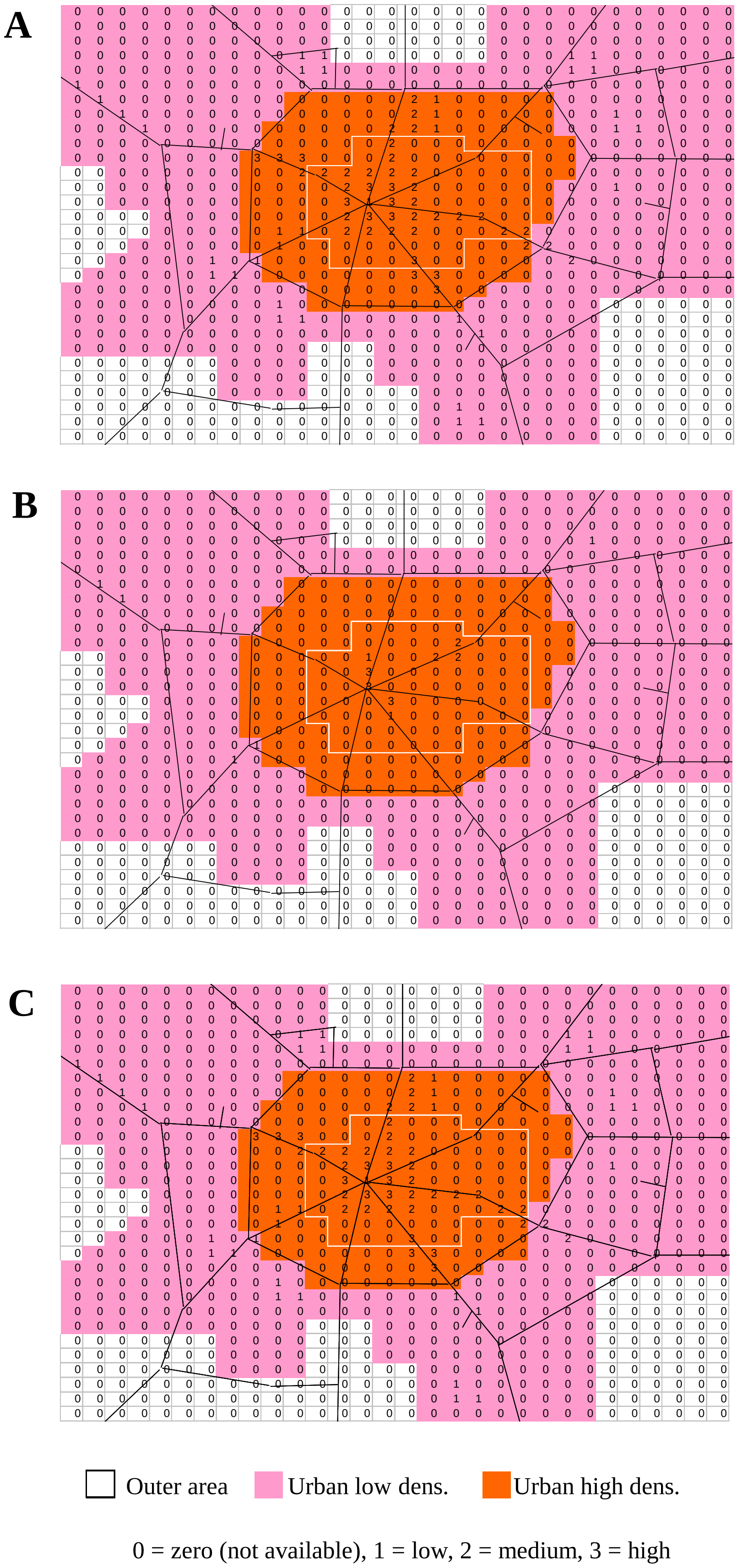


Figure 3: Maps of facilities for shopping (A), leisure (B) and recreation (C)

3.2 Results

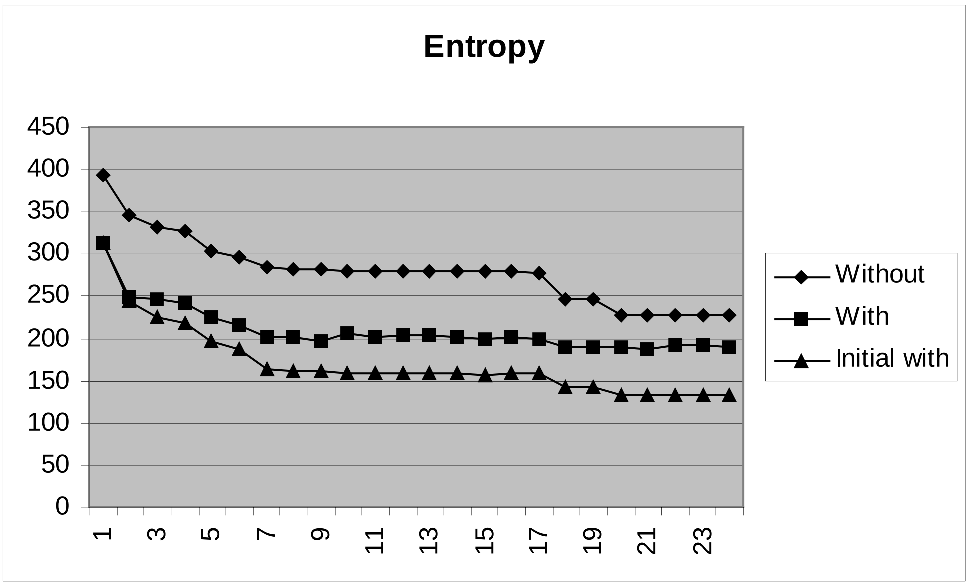
To measure the knowledge content of the mental map at any moment in time, an entropy measure and likelihood measure are defined as follows:

$$H(X_k) = -\sum_l \sum_s P(x_{lks}) \times \log_2 \{P(x_{lks})\} \quad (11)$$

$$L(X_k) = \sum_l P(x_{lk}^*) \quad (12)$$

where $H(X_k)$ is the total entropy of variable X_k across locations, $L(X_k)$ is the total likelihood of correct beliefs about X_k across locations, $P(x_{lks})$ is the belief in value x_{lks} of X_k in cell l and $P(x_{lk}^*)$ is the belief in the true value. Thus, where the entropy expresses the degree of uncertainty or information content, the likelihood indicates the accuracy of the knowledge.

The graphs in Figure 4 represent the entropy and likelihood for the availability of shopping facilities (as an example) as a function of time for three runs of the model. The runs differ with respect to the role of the belief network. In the first and third run (Without and Initial With), the network was excluded, so that beliefs are updated based only on observations made on the variable under concern. In the second run (With), the network was included to spread evidence among related variables. In the Initial With case, the network is used only to derive the initial beliefs (i.e., a-priori beliefs) and is further excluded in belief updating. In all three cases, the entropy decreases and likelihood increases monotonously indicating that the knowledge increases over time. Flat regions in the curves correspond to the periods in which only work activities are conducted and, hence, the same trips are repeated multiple times. The state of perfect knowledge would correspond with a zero value for the entropy and a value of 32 × 32 = 1024 for the likelihood measure. The fact that these levels are not attained means that the mental map still reflects incomplete knowledge of the area after all activities have been conducted.



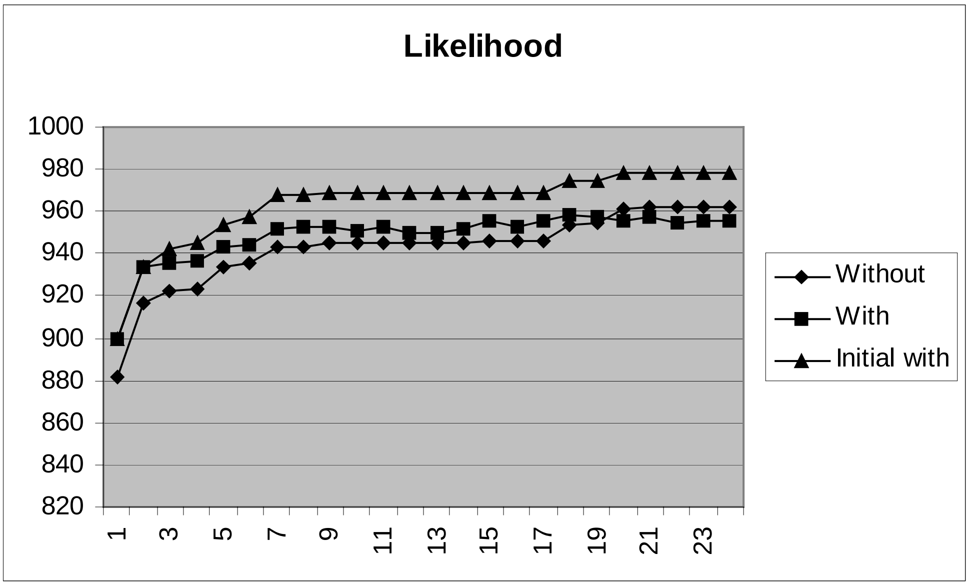


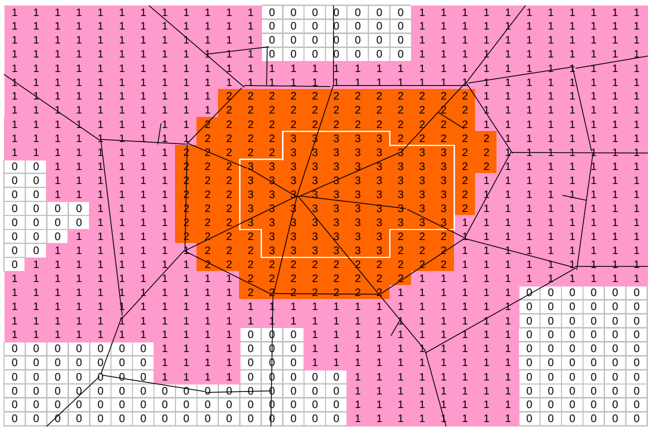
Figure 4: Entropy and likelihood of the mental map as a function of time (i.e., number of activities conducted)

The differences between the curves indicate the impact of the belief network on learning. Both in terms of the entropy and likelihood, the With condition improves the learning end result compared to the Without condition indicating that reasoning reduces uncertainty and increases the accuracy of the mental map. However, this effect of the belief network is to be attributed to a better starting point rather than to a higher learning rate. In effect, learning is even somewhat slower under the With condition. This is proven by the 'Initial with' graph, which combines the better starting position with the higher learning rate. Similar patterns emerge for the other variables as well (not shown).

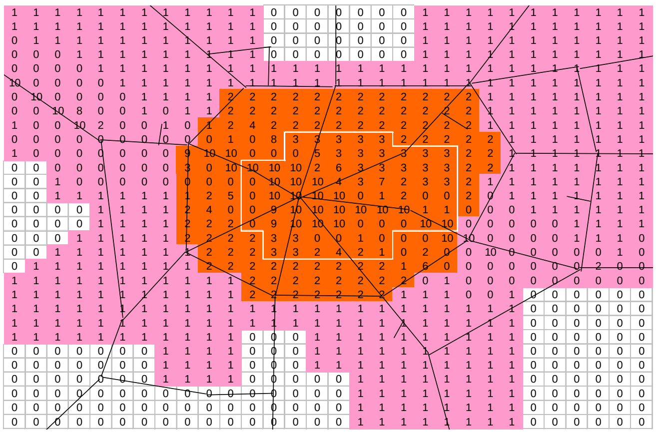
This finding suggests that the best learning result is obtained when generic knowledge is used to derive a-priori beliefs, in an initial state, and is not used afterwards in belief updating in response to observations. A closer look at differences in state of the mental map under the different conditions reveals the impact of spatial reasoning. When spatial reasoning is involved, the beliefs display a tendency to regress to theoretically expected values. This follows from the fact that indirect evidence, which has a theoretical component, is combined with direct evidence each time a belief is updated. Theory has the advantage that the system is less sensitive to biases or lack of (sensitivity of) observations on a certain attribute. However, in the case when all attributes are observed with identical outcome-probability functions, as in the simulations here, spatial reasoning is counter-productive in the sense that it introduces conceptual biases. From a modeling point of view the behavior may however be adequate as real mental maps may display similar biases.

To illustrate the state transitions of the mental map underlying these measures, Figure 5 represents the mental map under the With condition at three time moments in time again regarding the shopping-facility-availability variable. The numbers represent rounded probabilities on a 0 – 10 scale (0 is certainly no, 10 is certainly yes and 5 means maximum uncertainty). Map A shows the a-priori beliefs in the initial state. The initial probabilities differ only between the areas distinguished by the urban-area-type variable, as this is the only information initially available. For example, in the inner city the a-priori probability of a shopping facility is higher than in high-density area around the center. In the outer area the probability is even considered to be zero. Map B depicts the state of the mental map after the first activity has been conducted. This activity is a work activity so that the route of the trip

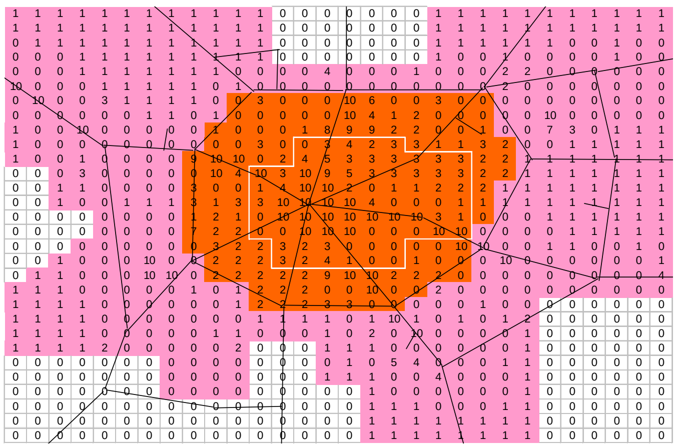
A



B



C



0: 0%, 1: 10 %, …, 10: 100%

Figure 5: Predicted beliefs of availability of shopping facilities *with network* ($\theta = 4 - 1.25\ d$, no knowledge decay). A: after 1 activity; B: after 23 activities.

describes the shortest path between the home and work location (see Figure 2). The simulated effects of observations made during the trip are clearly visible as a band along both sides of the route of the trip of approximately 2-3 cells wide. Within this band the probabilities tend to be approximately equal to 0% for locations were there are no shops and 100% for locations were there are shops (cf. Map A in Figure 3) meaning that uncertainty has been reduced almost completely. The width of this band reflects the assumed distance-decay function of the observation sensitivity. However, the effects of observations are not completely homogeneous within the band. For some of the cells there is more uncertainty left as a consequence of conflicting outcomes of observations (due to limited sensitivity). The probability of conflicting outcomes increases and their impact on beliefs decreases with increasing distance from the route.

Map C portrays the state of the mental map of shopping facilities after all 23 activities have been conducted. The uncertainty about where shopping facilities are located has decreased considerably as indicated by the increasing number of zero and ten-value cells. However, the reduction is not homogeneous across locations, but rather reveal the routes that the individual frequently takes to conduct the activities at different locations in the area. The center area and areas around frequently traveled routes are well known, while other parts of the city are still largely or completely unknown. As it turns out, the incremental learning over time has an effect also on the location choice of shopping activities (not shown). For example, at first the individual chooses the nearest location from home being unaware of the relative attractiveness of the locations observed on routes. It appears that this location has a low attractiveness and the next time the individual tries the second nearest location from home. After experiencing that also this location has a low attractiveness, the next shopping trip switches back again to the nearest location. Because of the distance, the expected utility of the other shopping locations known to the individual is lower, even though there is full uncertainty about the attractiveness of these alternatives.

4. Conclusions and discussion

In this paper, we have put forward a Bayesian belief network for modeling the dynamics of mental maps. In the present paper, we have restricted the model to the problem which destination location and their attributes are known to an individual traveler and how such maps evolve over time as a function of the implementation of activity-travel schedules. The potential of the approach has been illustrated using numerical simulation. The results of the simulation demonstrate the potential of the suggested approach.

When evaluating the results of the model, it should be realized that only part of the full problem of mental maps of urban environments have been covered. Four additional lines of research are required to develop a full-fledged model. First, the model presented in this paper dealt with the perception of locations. In addition, a model how individuals learn about the road network structure needs to be developed to complement the present model. Secondly, although we articulated its importance, the present model has focused on how people learn about their environment as a result of conducting a particular set of activities. However, people may also learn from other sources such as news media and social contacts, implying that either the present model should be further elaborated along for this case, or that another model should be developed to simulate the effect of sources other than travel on the dynamics of mental maps. Finally, we have only considered the perception and cognition dimension. In conducting their activities, people will also experience their environment, thereby forming and updating judgments. Again, the development of a model how individuals form and update their value judgments about the attributes of the choice alternatives in their mental maps constitutes another critical component of a realistic dynamic micro-simulation system of activity-travel behavior.

Furthermore, the model of the mental map and cognitive learning includes several sets of parameters that need to be estimated on data. These include the parameters of the observation-outcome probability function and the conditional probabilities associated with nodes in the belief network. Possibly, data about how individuals perceive and update their beliefs could be collected in virtual reality environments where the experimenter is able to control the physical aspects of the environment and subjects can implement their trips and activities in a simulated environment. Alternatively, the model may be calibrated on revealed knowledge patterns stratified by level of expertise and activity-pattern characteristics of a larger sample of individuals. Much additional work is needed to develop feasible and effective methods of data collection and estimation/calibration.

One immediate application of the suggested approach concerns the delineation of choice sets. Depending upon one’s definition of a choice set, the model presented in this paper could either be applied in a straightforward manner or could be the stepping stone of a model to delineate dynamic choice sets. If one would define a choice set as the set of choice alternatives known to an individual traveler, then the model presented in this paper could be used as is. If, on the other hand, one would entertain a stricter definition, arguing that an alternative would only belong to a choice set if it is seriously considered for choice, then some additional assumptions have to be made. We intend to address this problem in more detail in a forthcoming paper.

Ultimately, all these components will be added as agents to the Albatross system (Arentze and Timmermans, 2000). The agents dealing with the dynamics of mental maps as discussed here will then be used to simulate the dynamics of individuals' cognitive representation of space and transport networks, and to simulate the dynamics of their value judgment. The application of these agents in a micro-simulation will then generate at each point in time an individuals' mental map and subjective assessment of the available choice alternatives. The rules embedded in the current version of Albatross, representing choice heuristics, can then be used to schedule activities in time and space.

References

Arentze, T.A. and Timmermans, H.J.P.(2000), Albatross: *A Learning Based Transportation Oriented Simulation System*, European Institute of Retailing and Services Studies, Eindhoven

Arentze, T.A. and Timmermans, H.J.P. (2003), Modeling learning and adaptation processes in activity-travel choice: a framework and numerical experiments, *Transportation,* 30, 37-62.

Golledge, R.G. (1993), Geographical perspectives on spatial cognition, In, T. Gärling and R.G. Golledge (eds.), *Behavior and Environment*, North Holland, Amsterdam, pp. 16-46.

Heckerman, D., A. Mandani and M.P. Wellman (1995) Real-world Applications of Bayesian Networks, *Communications of the ACM*, **38**, 24-26.

Heijden, R.E.C.M. van der and Timmermans, H.J.P. (1984), Modelling choice set generating processes via stepwise logit regression procedures: some empirical results, *Environment and Planning A*, 16, 1249-1255.

Horton, F.E. and Reynolds, D.R. (1971), Effects of urban spatial structure on individual behavior, *Economic Geography*, 47, 36-48.

Neapolitan, R.E. (1990) *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*, John Wiley & Sons, New York.

Pearl, J. (1988) *Probabilistic Reasoning In Intelligent Systems*, Morgan Kaufmann, San Mateo, California.

Potter, R.B. (1979), Perception of urban retailing facilities: an analysis of consumer information fields, *Geografiska Annaler B*, 61, 33-38.

Russel, S. and P. Norvig (1995) *Artificial Intelligence: A Modern Approach*, Prentice Hall, Englewood Cliffs, NJ.

Smith, G.C. (1976), The spatial information fields of urban consumers, *Transactions of the British Institute of Geographers*, 1, 175-189.

Spiegelhalter, D.J., A. Philip, S.L. Lauritzen and R.G. Cowell (1993) Bayesian Analysis in Expert Systems, *Statistical Science*, **8**, 219-283.

Timmermans, H.J.P., Arentze, T.A. and Ettema, D. (2003), Learning and adaptation behaviour: empirical evidence and modelling issues, *Proceedings ITS Conference*, Eindhoven, April 1-4.

Timmermans, H.J.P., van der Heijden, R.E.C.M. and Westerveld, H. (1982), Perception of urban retailing environments: an empirical analysis of consumer information and usage fields, *Geoforum*, 13, 27-37.